

Das «Zündeln» im Urteil: Alterstrends und psychometrische Diagnostizierbarkeit der zivilrechtlichen Verantwortlichkeit nach § 828 BGB*

Judgments on fire-setting: Age trends and psychometrical assessments of liability (§ 828 BGB)

Wilfried Hommers**

Universität Würzburg

Zusammenfassung: Die Regelungen des § 828 BGB zur zivilrechtlichen Verantwortlichkeit Minderjähriger für unerlaubte Handlungen, z. B. Brandlegungen, konstatieren zwei Kompetenzen: Unrechtserkenntnis und Vergeltungspflichtverständnis. Im Sachverständigenbeweis kann ein Minderjähriger befreit werden von der Haftung. Die Rechtsprechung unterstellt: Unrechtserkenntnis enthebt des Nachweises des Vergeltungspflichtverständnisses. Daher ist, de lege ferenda, die Altersgrenze vom vollendeten siebenten Lebensjahr durch diesbezügliche Entwicklungsvorgänge zu stützen. Außerdem sind, de lege lata, geeignete Methoden für die individuelle Erfassung der Kompetenzen bereitzustellen. Mit Hilfe quantitativer Beurteilungen von faktoriell aus Schadenshöhe-, Verschulden- und Entschuldigung-Informationen zusammengesetzten Geschichten und von IQ-Subtests werden die Resultate einer empirischen Studie berichtet, die die impliziten und expliziten gesetzlichen Annahmen stützen und die psychometrisch fundierte Methoden zur individuellen Erfassung der beiden Kompetenzen bereitstellten. Fehlende Korrelationen zwischen IQ-Maßen und moralischen Urteilen zeigten, daß sich die Erfassung der erforderlichen Kompetenzen nicht durch IQ-Subtests ersetzen ließe.

Abstract: In German civil law, children's liability for arson, as for other torts, rests on the preassumption that the individual child has two competencies: knowledge of right and wrong when doing the harm and some understanding of the duty to make recompense for the harm. By an expertise proof of incompetency a minor may be exempted from damages. As a rule, existing right-wrong-knowledge proves the understanding of the repair-duty. De lege ferenda, the age limit of seven years ought to be based on at least some feature of the developments of the relevant concepts. De lege lata, methods for the assessment of an individual's competencies are needed. Using quantitative judgments on factorial (damage \times culpa \times apology information) fire-setting stories as well as other methods (e. g. IQ-subtests) the results of an original empirical study are reported which were (a) in support of the implicit and explicit legal preassumptions about minor's competencies relevant for their liability, and (b) provided psychometrically based tools for the individual diagnoses of the distinctive appreciations of the right-wrong-knowledge and of the understanding of the repair-duty. The lack of correlations of judgments and intelligence subtests showed that the capacity assessment cannot be replaced by intelligence testing.

Die Brandstiftung durch Kinder oder Jugendliche als ein Teil der Brandschäden überhaupt ist ein entwicklungsbedingtes Alltagsproblem mit verschiedenen Seiten. Die Prävalenz von Brandverursachungen durch Kinder oder Jugendliche in der Bundesrepublik Deutschland wurde mit 7000 bis 8000 pro Jahr angegeben (DeBoor, 1955), aus der polizeilichen Kriminalstatistik

für 1960 gingen 2080 Fälle fahrlässiger und vorsätzlicher Brandstiftung durch Kinder (bis auf 193 Fälle durch Jungen verübt) hervor (Dauner, 1980, S. 13). Das Statistische Jahrbuch 1986 (S. 337) wies für 1984 185 wegen fahrlässiger oder vorsätzlicher Brandstiftung (§§ 306–309 StGB) verurteilte Jugendliche aus; 1187 aufgeklärte bei 9.698 bekanntgewordenen vorsätzlichen Brandstiftungen und 421 aufgeklärte bei 8.758 bekanntgewordenen fahrlässigen Brandstiftungen waren auf Strafunmündige zurückzuführen (S. 335), bei ca. 4 Millionen bekanntgewordenen Straftaten überhaupt. Das Interesse und Spielen mit Feuer führt auch vor einem polizeilich bekanntgewordenen Delikt gelegentlich nicht nur zur Wahrnehmung eines unge-

* Widmung: Prof. Dr. phil. Dr. med. Hermann Wegener zum 70. Geburtstag.

** Danksagung: Mit Unterstützung durch die DFG: Ho 920/2–2. Herr Dipl.-Psych. K. Feld führte die MANOVA-Auswertungen am Rechenzentrum der Universität Würzburg aus. Frau M. Pirkner fertigte das Manuskript, die Tabellen und die Abbildung.

wöhnlichen Umgangs mit Feuer als Verhaltensproblem eines Kindes oder Jugendlichen, sondern auch zur Vorstellung von in der Mehrzahl männlichen Kindern oder Jugendlichen in der kinderpsychologischen oder kinderpsychiatrischen Erziehungsberatung (vgl. die US-amerikanische epidemiologische Studie von Achenbach & Edelbrock, 1981, S. 26 und 38). Von den dadurch ausgelösten präventiven, erzieherischen (Ell, 1983) oder therapeutischen (Holland, 1969) Maßnahmen zur Vermeidung von Brandverursachungen durch Kinder zu unterscheiden sind die diagnostischen Interessen der Psychologie an der Beurteilung des Reifestandes zur zivil- oder strafrechtlichen Verantwortlichkeit nach § 828 BGB (Bürgerliches Gesetzbuch) oder § 3 JGG (Jugendgerichtsgesetz). Derartige Begutachtungsaufträge kommen zu den sonst seltenen Deliktsfähigkeitsbeurteilungen nach § 828 (2) BGB sogar relativ häufig vor. Ell (1983) weist z. B. auf 35 von ihm begutachtete Fälle hin. Das Vorkommen solcher Begutachtungsaufträge ermöglichte auch klinische Vergleichsstudien unter Fall-typisierenden Fragestellungen. Dauner (1980) verglich z. B. 30 Brandstiftung-Begutachtungen von Kindern und Jugendlichen mit 30 Diebstahl-Begutachtungen dieser Altersgruppen und fand, daß das «Zündeln» von Kindern weitaus weniger als der Diebstahl Ausdruck kindlicher Neurosen oder eines frühkindlichen exogenen Psychosyndroms sei (S. 131 f.). Die «Brandstifter» erschienen ängstlich-empfindsamer, trotzig-aggressiver, kontaktgestörter als die «Diebe» (S. 135). Klosinski (1985) verglich unter § 3 JGG je 10 zu begutachtende jugendliche, männliche Brandstifter und Sexualdelinquenten. Erstere hatten häufiger cerebral-organische Funktionsstörungen und eine tatrelevante «Vater-Sohn-Problematik», während letztere nicht unterbegabt waren und eine «Mutter-Sohn-Beziehungsstörung» bei Vaterabwesenheit hatten. Bei beiden Autoren findet man auch weitere Angaben über ältere deutsch- oder englischsprachige Literatur (außerdem Canter, 1980, und Prins, 1986).

1. Fragestellung

Die folgend berichtete Studie beschränkt sich auf die kognitiv-psychologischen Kriterien der zivilrechtlichen Verantwortlichkeit nach § 828

BGB (Undeutsch, 1967) und will sowohl zu der Frage der Entwicklungstrends als auch zur Frage der individuellen Erfäßbarkeit der betreffenden Kriterien durch Urteilsmethoden einen Beitrag leisten. Zu den hier bestehenden Fragestellungen *de lege lata* (forensisch-psychologische Beurteilung eines Minderjährigen) und *de lege ferenda* (empirische Begründung der Altersgrenze) stellte Hommers (1983, 1989) empirische Befunde und Methoden dar, die aus der kognitiven Entwicklungspsychologie vorlagen, ohne auf diese rechtspsychologische Frage abgezielt zu haben. Der vorliegende Beitrag befaßte sich daher spezifisch mit diesen Fragen und wählte wegen seines nicht unerheblichen Vorkommens als geeigneten Untersuchungsinhalt die Brandstiftung durch Kinder.

Die Frage des Entwicklungstrends zwischen Vor- und Grundschulalter ist *de lege ferenda* unter der allgemeinen Entwicklungsannahme des § 828 (1) BGB von Interesse, wonach mit der Vollendung des siebenten Lebensjahres die zivilrechtliche Verantwortlichkeit nicht mehr ausgeschlossen ist, was angesichts der Beweislast seitens eines älteren minderjährigen Schädigers dahin auszulegen ist, daß die zivilrechtliche Verantwortlichkeit von da ab als in der Regel erreicht angenommen wird. Ergebnisse von Begutachtungen über minderjährige Brandstifter und Diebe wurden von Dauner (1980) zum Anlaß genommen, zu der deliktsrechtlichen Altersgrenze vom vollendeten siebenten Lebensjahr Stellung zu nehmen. Da die Regelung der Altersgrenze aber nicht so sehr von Begutachtungsergebnissen wie von der allgemeinen Entwicklung der relevanten Kriterien abhängen muß, erschien die diesbezügliche Untersuchung von Nichttätern erforderlich, wobei die Untersuchung von dort gar nicht beachteten Vorschülern eingeschlossen werden konnte, die ja niemals einer zivilrechtlichen Begutachtung zugeführt würden.

Die Frage der individuellen Erfäßbarkeit der einschlägigen Kriterien ist offensichtlich *de lege lata* durch § 828 (2) BGB von Interesse, wonach es möglich ist, daß ein Minderjähriger bis zur Vollendung des 18. Lebensjahres zivilrechtlich nicht verantwortlich ist, wenn er nachweisen kann, daß er die erforderliche Einsicht nicht besaß. Die Kriterien der zivilrechtlichen Verantwortlichkeit sind seit einem Reichsgerichtsurteil

von 1902 grundsätzlich das Vergeltungspflichtverständnis, das Unrechtsbewußtsein gegenüber dem Mitmenschen und das Sorgfaltpflichtverständnis (vgl. Undeutsch, 1967), wenngleich ersteres in der Regel als gegeben angesehen wird, wenn eines der letzteren beiden ausgebildet ist.

Die Erhebungsmethode für die Kriterien der zivilrechtlichen Verantwortlichkeit war wegen der angestrebten individuellen Erfäßbarkeit von zentralem Interesse für die Studie. Schon das Reichsgericht (vgl. Hommers, 1983) stellte nur geringe Anforderungen an das Verständnis der Vergeltungspflicht («irgendein Verständnis»). Demgegenüber zielen Gutachter auf adäquate Äußerungen in mehr oder weniger gezielten Befragungen ab. Nach Dauner (1980, S. 122) gaben immerhin 23% der 30 von ihr untersuchten Brandstifter im Alter zwischen 7 und 13 Jahren im Interview Absichten zur Wiedergutmachung an (ohne Alterseinfluß) gegenüber 3% bei der in der im Alter gleichartigen Vergleichsgruppe von Diebstahl-Begutachtungen. Solche Methoden der Gutachter sind möglicherweise schon methodisch (Produktion einer sprachlichen Äußerung versus Beurteilungen von audio-visuellen Informationen) zu anspruchsvoll oder treten wegen der Vorsätzlichkeit des Delikts (Diebstahl gegenüber zumeist fahrlässiger Brandstiftung) unterschiedlich wahrscheinlich auf (bei absichtlichen Taten wird man sich vermutlich eher nicht mit dem Gedanken an aktive Wiedergutmachungsabsichten befassen). Weiterhin ist die höchstrichterliche Anforderung an die Kriterien nicht nur gering, sondern die erforderliche Einsicht besteht aus zwei Komponenten. Neben der erkennenden Komponente der Einsichtsfähigkeit (Wissen-Können) ist die werten-de Komponente (Werten-Können) zu untersuchen. Gerade bei letzterer fehlte es aber bislang an geeigneten Methoden. Es blieb daher bisher ungeprüft, ob die häufig vorgenommene stellvertretende Reifestandserfassung mit Intelligenzleistungstests (z.B. dem HAWIK-Subtest Allgemeines Verständnis) generell ausreichen konnte (Bresser, 1972; Eisen, 1977, S. 320; Ell, 1983, S. 65; kritisch dazu: Hommers, 1983, S. 29–34). Denkbar wäre, daß zwar das Wissen-Können durch die IQ-Subtests erfaßt werden könnte, nicht aber das Werten-Können.

Als Alternative zur impliziten oder expliziten sprachlichen Kommunikation der geforderten

Wertungs-Kompetenzen während einer Exploration wird daher im folgenden die aus Urteilen ableitbare moralische Unterscheidungsfähigkeit eingesetzt, womit an das klassische Discernement-Konzept des rechtlichen Denkens angeschlossen wird (Hommers, 1983, S. 16). Diese Methode ermöglicht im Prinzip sowohl das Vergeltungspflichtverständnis als auch das Unrechtsbewußtsein bei gleichem Anforderungsniveau in der Äußerungsform zu bestimmen (und kann in folgenden Studien auch das Sorgfaltpflichtverständnis erfassen). Das niedrig ange-setzte Niveau kommt aber in der Untersuchung nicht nur in der Äußerungsform, sondern auch im kognitiven Gegenstand zum Tragen. Es geht um die Unterscheidung der Entschuldigung von dem Sich-Nicht-Entschuldigen, was im Vergleich zur Haftung im zivilrechtlichen Sinne (voller materieller Schadensersatz) inadäquat erscheinen mag, aber von der geringen Anforderung der Formulierung «irgendein Verständnis» adäquat. Hinsichtlich des Vergeltungspflichtverständnisses war aber als Konsequenz dieser Variablenwahl von besonderem Interesse, ob auch weitere Unterscheidungsleistungen in Abhängigkeit vom Alter feststellbar waren. Daher wurde auch das Reparierenhelfen durch das Kind als Bedingung aufgenommen. Diese Bedingung ist im Sinne der Equity-Theorie (Walster, Walster & Berscheid, 1978) an den demnach ubiquitär als adäquat bewerteten exakten Ersatz angeglichen. Durch die Vorgehensweise wurden also die Unterscheidungsleistungen verschiedener Informationenteile (Schadenshöhe, Verschulden) mit den Wiedergutmachung-Informationen verglichen und zu zwei Testzeitpunkten in unterschiedlicher Aufgabenschwierigkeit (Instruktion mit gegenübergestellten ein-faktoriellen Stimuli versus Hauptphase mit kompletten, unregelmäßig angeordneten dreifaktoriellen Stimuli) erhoben. Schließlich war es geboten, die Beziehungen der hiermit erlangten Befunde zu Maßen der allgemeinen intellektuellen Entwicklung zu bestimmen, die als Indikatoren des Wissen-Könnens gelten können (z.B. der Untertest «Allgemeines Verständnis» im HAWIK-R, Tewes, 1983).

Die individuelle Erfäßbarkeit der Kriterien warf aber auch statistisch-methodische Probleme auf. Daher ergab sich die aufgrund einer kritischen Stimme gegenüber einer statistische und

mathematisierend-experimentelle Methoden verwendenden Forensischen bzw. Rechts-Psychologie (Bresser, 1988) günstige Gelegenheit, konstruktive Beiträge einer derartigen Psychologie auch in diesem Gebiet der Psychodiagnostik aufzuweisen.

2. Methode

Gut-Böse-Urteile von Vorschülern ($N = 33$), Grundschulern ($N = 38$), Hauptschülern ($N = 37$) und Erwachsenen ($N = 40$) wurden zu zwölf bebilderten Geschichten über Brände erhoben, die durch ein Kind auf einem Bauernhof verursacht wurden. Die Geschichten waren aus drei Nachtatverhalten- (Reparaturhilfe, Entschuldigung, keine Entschuldigung), zwei Verschulden- (Versehen, Wut) und zwei Schaden-Bedingungen (Hundehütte, Scheunentor) faktoriell kombiniert (vgl. Abbildung 1). Für die Darstellung der Verschulden-Stufen wurden zwei Bilder verwendet, für die Stufen der beiden anderen Faktoren nur je eins.

Zur Einübung in die Aufgabe wurden in der Instruktionsphase zuerst die Geschichtenbedingungen jedes Faktors gegenüberstellend audiovisuell dargeboten (in der Reihenfolge Schaden, Nachtatverhalten, Verschulden) und einzeln beurteilt, dann waren vier komplette Geschichten in Form des Bilderordnen-Tests (BO im HAWIK-R; Tewes, 1983) in die richtige Abfolge zu legen und danach zu beurteilen (die Hälfte der Kinder fuhr in dieser Weise in der Hauptphase fort, die andere Hälfte bekam in der Hauptphase irreguläre Anordnungen der Geschichten simultan angeboten, was nach der statistischen Prüfung keine Methodeneffekte erzeugte) und schließlich waren 3 bzw. 2 weitere Geschichten unmittelbar vor der Hauptphase der Beurteilung der 12 faktoriellen Kombinationen ohne weitere Hilfen oder Korrekturen des Versuchsleiters zu beurteilen. Die Skala bestand aus 13 weißen und 13 schwarzen von der Mitte her nach den Seiten hin zunehmenden, bis zu 6 cm langen und jeweils .5 cm breiten Balken zur Veranschaulichung des Urteilsinhalts und seines Ausmaßes. Es erfolgte keine Endanker-Definition durch vollständige Beispiel-Geschichten.

Die Urteile über die faktoriell kombinierten Stimuli dienten einerseits zur Bestimmung der Reliabilität der individuellen Effektstärken, die

als lokale (bedingte) Effekte durch Aufsummierungen gebildet werden konnten. Die Reliabilität wurde durch Cronbachs α für die Differenzsummen und durch Meßwiederholungskorrelationen der Urteile unmittelbar vor der Hauptphase und den dazu im Inhalt gleichen Stimuli in der Hauptphase geschätzt. Auf der Basis der Reliabilitätsschätzungen sollten dann die Kriterien individuell aus den individuellen Effektstärken «diagnostiziert» werden. Die Urteile der Instruktionsphase boten einen Anhaltspunkt dafür, was sich spontan, aber unter erleichterten Bedingungen ergeben würde. Andererseits konnten die Alterstrends der Urteile auch hier geprüft werden. Dabei ging es auch darum, ob die Wirkung der Entschuldigung auf moralische Urteile von Kindern und Erwachsenen gegenüber den Wirkungen von Verschulden und Schaden vorherrscht und mit dem Alter stabil bleibt, was aus früheren Befunden über die Schadenswiedergutmachung bzw. ihrer Komponenten, Entschuldigung und Entschädigung, mit Darstellungen anderer Delikte (Briefmarkenruinierungen, Diebstahl) zu erwarten war (Hommers, 1988 a, b).

Außer den Urteilen wurden von den nicht erwachsenen Probanden die Untertests MT, AV und WT des HAWIK-R (Tewes, 1983) erhoben, um zu prüfen, ob sich die individuellen Effektstärkenmaße aus den Maßen des intellektuellen Entwicklungsstandes hinreichend vorhersagen lassen oder nicht und ob sie deswegen in der forensischen Begutachtung der zivilrechtlichen Verantwortlichkeit als Kriterien der wertenden Komponente des Vergeltungspflichtverständnisses und des Unrechtsbewußtseins besonders zu berücksichtigen sind oder nicht. Die Untersuchung begann mit dem MT. Der WT wurde zwischen die Instruktionsphase und die Hauptphase der Geschichtenbeurteilungen gelegt. Der AV erfolgte am Ende der Untersuchung. Die Untersuchung dauerte insgesamt bis zu 45 Minuten und wurde von mehreren trainierten, weiblichen Testleitern durchgeführt.

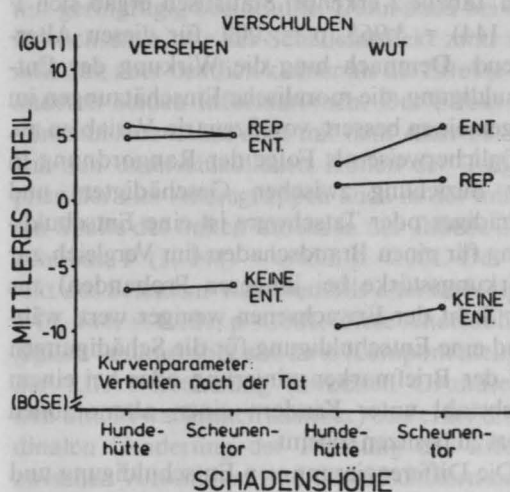
3. Ergebnisse

3.1 Alterstrends in der Beachtung der Stimulusinformationen

Ein genereller Haupteffekt des Alters trat nicht auf, $F(3,144) = .331$, $p = .803$. Die Geschich-

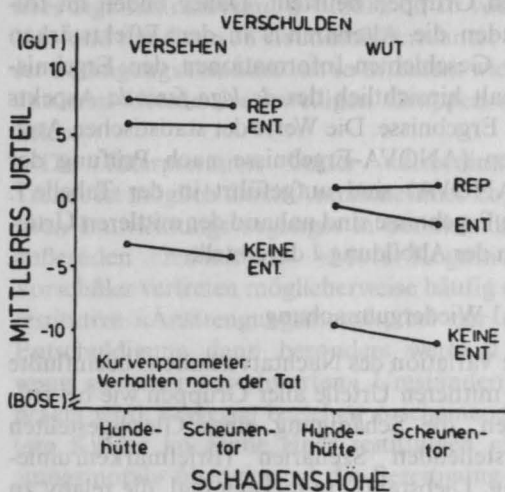
VORSCHÜLER

(5-6 Jahre, $N=33$, $73 \leq S_E \leq 169$)



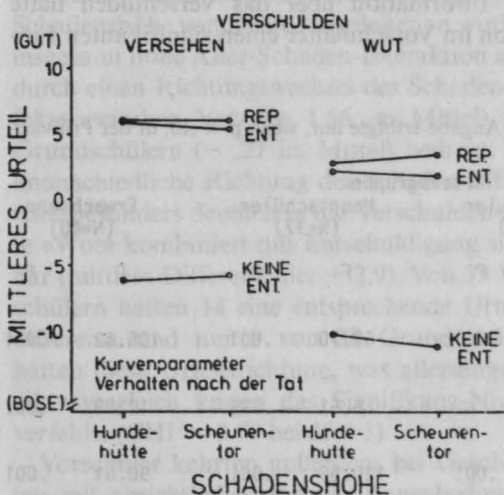
HAUPTSCHÜLER

(11-16 Jahre, $N=37$, $47 \leq S_E \leq 111$)



GRUNDSCHÜLER

(7-10 Jahre, $N=38$, $55 \leq S_E \leq 121$)



ERWACHSENE

(18-35 Jahre, $N=40$, $51 \leq S_E \leq 99$)

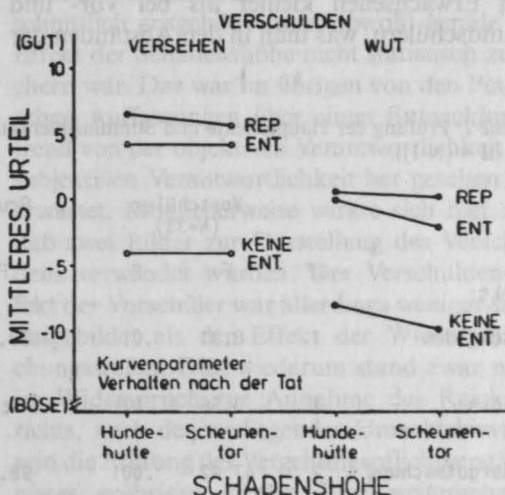


Abbildung 1: Mittlere Urteile von 4 Altersgruppen über 12 Bild-Geschichten, die zwei Brandschäden (horizontale Achse), das Nachtatverhalten (Kurvenparameter) und das Verschulden (linker und rechter Graph jeder Gruppe) faktoriell kombinierten. Man sieht den mit dem Alter zunehmenden Verschulden-Effekt im Vergleich der Niveaus der linken und rechten Teilgraphen, den mit dem Alter teils abnehmenden, teils zunehmenden Effekt des Nachtatverhaltens in den Kurvenabständen und den altersabhängigen Schaden-Effekt in den Kurvensteigungen.

tenkombinationen über Brandstiftungen und Nachtatverhalten wurden also gleich streng von allen Gruppen beurteilt. Daher bilden im folgenden die Alterstrends in den Effektstärken der Geschichten-Informationen den Ergebnishalt hinsichtlich des *de lege ferenda* Aspekts der Ergebnisse. Die Werte der statistischen Analysen (ANOVA-Ergebnisse nach Prüfung der MANOVA) sind aufgeführt in der Tabelle 1. Die Ergebnisse sind anhand der mittleren Urteile in der Abbildung 1 dargestellt.

3.1.1 Wiedergutmachung

Die Variation des Nachtatverhaltens beeinflusste die mittleren Urteile aller Gruppen wie bei früheren, die Schädigung eines Gleichgestellten darstellenden Szenarien (Briefmarkenruinierung, Diebstahl) am meisten (vgl. die relativ zu den Kurvensteigungen und den Teilgraphen-Niveaus großen Kurvenabstände in Abbildung 1). Statistisch hing der totale Effekt der drei Wiedergutmachungsstufen aber vom Alter ab: $F(6, 280) = 3.32, p < .02$ unter der Greenhouse-Geisser-Korrektur.

Im einzelnen war der Einfluß von Entschuldigung im Unterschied zu eigenen früheren Ergebnissen mit einem Briefmarkenruinierungs- und einem Süßigkeiten-Diebstahl-Szenario (Hommers, 1988a, b) bei Hauptschülern und Erwachsenen kleiner als bei Vor- und Grundschulern, was man in den Abständen der

entsprechenden Kurven in den vier Graphen und in der mittleren Spalte des mittleren Blocks von Tabelle 2 erkennt. Statistisch ergab sich $F(3, 144) = 3.965, p = .009$, für diesen Alterstrend. Demnach hing die Wirkung der Entschuldigung, die moralische Einschätzungen im allgemeinen bessert, von Szenario-Variablen ab. Möglicherweise als Folge der Rangordnung in der Beziehung zwischen Geschädigtem und Schädiger oder Tatschwere ist eine Entschuldigung für einen Brandschaden (im Vergleich zur Wirkungsstärke bei jüngeren Probanden) aus der Sicht der Erwachsenen weniger wert, während eine Entschuldigung für die Schädigungen bei der Briefmarkenruinierung oder bei einem Diebstahl unter Kindern einen altersstabilen Wert zu besitzen scheint.

Die Differenzierung von Entschuldigung und Reparierenhelfen nahm dagegen mit dem Alter zu, da Hauptschüler und Erwachsene die Geschichten mit der Bedingung Reparaturhilfe deutlich besser beurteilten als die Geschichten mit der Bedingung des Entschuldigens, $F(3, 144) = 9.854, p < .001$. Das geht sowohl aus den Abständen der entsprechenden Kurven in den vier Graphen als auch aus dem rechten Block der Tabelle 2 (mittlere Spalte) hervor.

3.1.2 Verschulden und Schaden

Die Information über das Verschulden hatte schon im Vorschulalter einen signifikanten Ein-

Tabelle 1: Prüfung der Haupteffekte und Stimulusinteraktionen (Angabe erfolgte nur, wenn $p < .10$; df der Prüfvarianz: $df = (N-1)$).

Effekt:	df	Vorschüler (N=33)		Grundschüler (N=38)		Hauptschüler (N=37)		Erwachsene (N=40)	
		F	P	F	P	F	P	F	P
Verschulden	1	8.31	.01	30.27	.001	63.70	.001	105.62	.001
Schaden	1	3.56	.07	< 1		5.51	.025	12.88	.001
Wiedergutmachung	2	67.92	.001	98.08	.001	120.56	.001	90.61	.001
Schaden x Verschulden	1							6.99	.02
Verschulden x Schaden x Wiedergutmachung	2							4.75	.02

fluß, während die Schadenshöhe die mittleren Urteile aller Altersgruppen, wenn überhaupt, nur geringfügig beeinflusste. Denn auch bei den Erwachsenen war der Schadenseffekt zwar signifikant, aber deutlich kleiner als die Effekte der anderen beiden Informationen. Der Effekt des Verschuldens nahm aber mit dem Alter zu, wie aus den unterschiedlichen Höhen der Teilgraphen der vier Altersgruppen auch in der mittleren Spalte des linken Blocks in der Tabelle 2 erkennbar, $F(3, 144) = 6.452$, $p < .001$. Der Effekt des Schadens war ebenfalls altersabhängig, $F(3, 144) = 6.021$, $p < .001$. Diese Altersabhängigkeit bestand aber aus zwei Komponenten, einer Effektverstärkung (zwischen Grundschulern und Erwachsenen sichtbar) und einer disordinalen (Änderung der Richtung des Effekts zwischen Vorschülern und Grundschulern sichtbar).

3.1.3 Disordinaler Entwicklungsbefund

Neben den zuvor genannten «quantitativen» (d.h. Effektstärken-bezogenen oder «ordinalen») Alterstrends im Verschulden-Effekt und in der unterschiedlichen Bewertung von Entschuldigen und Nicht-Entschuldigen ist auf zwei «qualitative» (d.h. nicht auf Effektstärken bezogene, «disordinale») Alterstrends einzugehen. Neben der zunehmenden Berücksichtigung der Schadenshöhe vom Grundschulalter an wird die insgesamt hohe Alter-Schaden-Interaktion auch durch einen Richtungswechsel des Schaden-Effekts zwischen Vor- (+ 1.56 im Mittel) und Grundschulern (- .27 im Mittel) bedingt. Die unterschiedliche Richtung des Schaden-Effekts wird besonders deutlich in der Verschulden-Stufe «Wut» kombiniert mit Entschuldigung sichtbar (mittlere Differenz hier + 2.9). Von 33 Vorschülern hatten 14 eine entsprechende Urteilsdifferenz, und nur 8 von 38 Grundschulern hatten diese Urteilsrichtung, was allerdings im Altersvergleich knapp das Signifikanz-Niveau verfehlte ($\chi^2 = 3.77$ bei $df = 1$).

Vorschüler kehrten außerdem bei Geschichten mit absichtlichen Brandstiftungsbedingungen die Ordnung von Entschuldigen und Reparaturhilfe um, so daß erstere von ihnen im Durchschnitt besser beurteilt wurde. Der Altersunterschied zwischen Vor- und Grundschulern erreichte diesbezüglich (jedoch nur) bei der

Kombination «Wut und Scheunentor» mit $\chi^2 = 10.99$ ein deutliches deskriptiv-exploratives Signifikanzniveau ($p < .001$), da 15 von 33 Vor- und nur 4 von 38 Grundschulern unter dieser Bedingungskombination so urteilten, wie die mittleren Urteile der jeweiligen Gruppen ausfielen.

Die Interpretation beider «disordinaler» Trends ist möglich durch Annahme eines kognitiven Entwicklungsvorganges in der sich darin äußernden restitutiven «Moral-Kognition». Vorschüler vertreten möglicherweise häufig eine restitutive «Anstrengungsmoral», bei der eine Entschuldigung dann besonders wertvoll ist, wenn sie unter «erschweren» Umständen erbracht wird. Dagegen beziehen anscheinend ältere Kinder im Sinne einer restitutiven «Leistungsmoral» überwiegend die Beseitigung des Schadens in die Urteile ein.

3.1.4 Fazit

Hinsichtlich der Entwicklungsannahme des § 828 (1) BGB erscheint der «disordinale» Befund besonders wertvoll, weil hier ähnlich wie bei Hommers (1986 a, b) ein Entwicklungstrend in der Schadenswiedergutmachung gefunden wurde, der die Altersgrenze (grob) bestätigte. In dieser Hinsicht war auch wichtig, daß der Verschulden-Effekt schon bei Vorschülern durchschnittlich ausgebildet war, obwohl gerade der Effekt der Schadenshöhe nicht statistisch zu sichern war. Das war im übrigen von den Piaget-schen Auffassungen über einen Entwicklungstrend von der objektiven Verantwortlichkeit zur subjektiven Verantwortlichkeit her gesehen unerwartet. Möglicherweise wirkte sich hier aus, daß zwei Bilder zur Darstellung des Verschuldens verwendet wurden. Der Verschulden-Effekt der Vorschüler war allerdings weniger stark ausgebildet als der Effekt der Wiedergutmachungsstufen. Dies wiederum stand zwar nicht im Widerspruch zur Annahme des Reichsgerichts, nach der vorliegendes Unrechtsbewußtsein die Prüfung des Vergeltungspflichtverständnisses erübrigen könne. Eine weitergehende Prüfung folgt jedoch noch. In diesem Szenario ergab sich also auch bei Veränderung einiger Aspekte insgesamt keine Widerlegung der expliziten und impliziten Entwicklungsannahmen des § 828 (1) BGB.

3.2 Individuelle Erfassbarkeit der Kriterien für § 828 (2) BGB

3.2.1 Hauptphase

Die internen Konsistenzen nach Cronbachs α , die aus den lokalen Effektstärken «Entschuldigung minus Nicht-Entschuldigung» und «Wut minus Versehen» (nicht jedoch beim lokalen Effekt von «Reparaturhilfe minus Entschuldigung») berechnet wurden, lagen mit $r = .80$ im Bereich der bei psychometrischen Tests erreichten Werte (vgl. Tabelle 2). Sie waren bei Vorschülern etwas geringer als bei den drei anderen Gruppen, was sich auch in den Meßwiederholungskorrelationen über eine Geschichtenkombination ergab. Das war besonders unter zwei Gesichtspunkten bemerkenswert. Einerseits handelte es sich um Differenzmaße, für die eine geringe Reliabilität zu erwarten war. Andererseits beruhten die Reliabilitätsschätzungen auf nur vier bzw. sechs Items. Ein angemessener Vergleich mit psychometrischen Tests würde also nach einer Spearman-Brown-Korrektur der Differenzreliabilitäten noch günstiger für die Eintragungen in Tabelle 2 ausfallen. Im fol-

genden wurden aber die empirisch ermittelten Werte zugrundegelegt.

Auf der durch die Reliabilitätsschätzungen gewonnenen psychometrischen Grundlage kann die Indikationstheorie der höchstrichterlichen Rechtsprechung zu § 828 (2) BGB geprüft werden. Diese besagte, daß vorhandenes Unrechtsbewußtsein die Annahme des Vorliegens *irgend-eines Vergeltungspflichtverständnisses* rechtfertigt. Für die individuellen Diagnosen *Unrechtsbewußtsein* und *Vergeltungspflichtverständnis* wurden dazu aus Vergleichszwecken die beiden statistisch einseitig definierten Kriterien von 5% und 20% zur Anwendung von α und SD der Altersgruppen gewählt. Die gekreuzten Diagnose-Häufigkeiten sind in der Tabelle 3 aufgeführt. Die jeweils unten links aufgeführten Häufigkeiten sind die Widerlegungsfälle ausgebildeten Unrechtsbewußtseins bei fehlendem Vergeltungspflichtverständnis. Demnach kam, sieht man einmal von den zumindest zivilrechtlich uninteressanten Erwachsenen ab, nur eine Widerlegung bei den Vorschülern vor, bei 41 bzw. 58 (20%-Niveau) Stützungen der Indikationstheorie. Gemessen an den Häufigkeiten der

Tabelle 2: Interne Konsistenzen nach Cronbach (α), Mittelwerte (M) und Standardabweichungen (SD) der gemittelten lokalen (bedingten) Effekte der Stimulusvariationen (W: Wut, V: Versehen, E: Entschuldigung, \bar{E} : Keine Entschuldigung, R: Reparaturhilfe; r_{tt} : Meßwiederholungsreliabilität der Urteile für Teilgruppen mit N = 19, 19, 18, 19, 75).

			Verschulden (W-V)			Entschuldigung (E- \bar{E})			Reparaturhilfe (R-E)		
			(6 "Items")			(4 "Items")			(4 "Items")		
	N	r_{tt}^1	α	M	SD	α	M	SD	α	M	SD
Vorschüler	33	(.64)	.50	-2.5	4.33	.78	12.6	7.65	-.35	-1.1	3.44
Grundschüler	38	(.95)	.78	-3.4	4.03	.94	12.4	7.67	-.01	0.7	2.39
Hauptschüler	37	(.95)	.88	-6.3	4.71	.88	9.4	5.26	.27	1.9	2.19
Erwachsene	40	(.92)	.81	-5.4	3.28	.91	8.2	5.81	.40	1.6	1.99
Insgesamt	148	(.89)	.77			.88			.11		

¹⁾ r_{tt} : Meßwiederholungsreliabilität der Urteile für Teilgruppen mit N = 19, 19, 18, 19, 75.

Tabelle 3: Gekreuzte Häufigkeiten für die psychometrisch auf dem 5%-Niveau gesicherten «Diagnosen» Unrechtsbewußtsein und Vergeltungspflichtverständnis aufgrund der individuellen Effektstärken (eingeklammerte Werte für das einseitige 20%-Niveau).

		Vorschüler		Grundschüler		Hauptschüler		Erwachsene	
		ja		ja		ja		ja	
		nein		nein		nein		nein	
Diagnose des Vergeltungspflichtverständnisses	ja	6(12)	19(17)	15(19)	20(19)	20(27)	14(10)	30(33)	5(4)
	nein	1(0)	7(4)	0(0)	3(0)	0(0)	3(0)	4(2)	1(1)

rechten Spalten (bedingte Grundraten) wichen diese Häufigkeiten mit $\text{CHI}^2 = 7.96$ bzw. 5.04 im Sinne der Hypothesenstützung ab. Die Häufigkeiten der Erwachsenen kann man als Folge der Anwendung der Gruppenparameter für Reliabilität und Standardabweichung ansehen, was einige sehr präzise urteilende Probanden möglicherweise schlechter darstellt als mit direkten Wiederholungsreliabilitäten.

Zur Frage der Wahl des geeigneten statistischen Entscheidungskriteriums könnte man von der Überlegung ausgehen, daß Erwachsene in der Regel (d. h. z. B. zu 95%) das sich in der Unterscheidungsfähigkeit ausdrückende Unrechtsbewußtsein besitzen. Bestimmte man davon ausgehend das Irrtumsniveau, ergaben sich jedoch bei den Hauptschülern und den Erwachsenen mittlere kritische Urteilsdifferenzen unter 1, was nicht sinnvoll erschien. Es wurde daher bestimmt, wie wahrscheinlich es ist, daß ein Proband jeder Altersgruppe zufällig einen mittleren Verschulden-Effekt oder Entschuldigung-Effekt größer/gleich 1 hat (also durchschnittlich *einen* Unterschied macht). So ergab sich, daß für 37, 30, 27 und 24% der Vor-, Grund-, Hauptschüler und Erwachsenen ein mittlerer Verschulden-Effekt größer/gleich 1 noch als zufällig zu werten war. Entsprechende Zahlen für die individuellen Entschuldigung-Effekte betrugen 39, 30, 29, 28%. Leitet man daraus ab, daß 24% das erreichbare (d. h. wegen der Erwachsenen) Sicherheitsniveau für die irrtümliche diagnostische Interpretation des *Durchschnittlich-Einen-Unterschied-Machens* als Unrechtsbewußtsein oder Vergeltungspflichtverständnis ist, dann müssen mindestens 2.14 bzw. 1.32 oder 1.14 für einen individuellen Verschulden-Effekt eines Vor-, Grund- oder Hauptschülers vorliegen und 2.51 bzw. 1.32 oder 1.28 für den individuellen Entschuldigung-Effekt, damit Unrechtsbewußtsein oder Vergeltungspflichtverständnis zugeschrieben werden. Da es sich um den Vorteil der Nicht-Haftung handelt, den man aufgrund der Zuschreibung von Unrechtsbewußtsein bzw. Vergeltungspflichtverständnis verliert, erscheint dieses Risiko-Niveau für die diagnostische Entscheidung angemessen. Über diesen altersspezifischen *cut-off*-Werten lagen im Verschulden-Effekt 16, 23, 33 und 36 Vor-, Grund-, Hauptschüler und Erwachsene bzw. im Entschuldigung-Effekt 29, 38, 36 und 37 Probanden.

3.2.2 Instruktionsphase

Schließlich kann man nach Vorliegen von Reliabilitätsschätzungen auch prüfen, wie gut sich aus der Instruktionsphase über das Unrechtsbewußtsein oder das Vergeltungspflichtverständnis psychometrisch-einzelfalldiagnostisch urteilen läßt. Die mittleren Urteile und die Standardabweichungen der individuellen Effekte sind in der Tabelle 4 aufgeführt. Durch Auszählung der vier kleinsten individuellen Effekte zeigte sich auch schon in der Instruktionsphase die psychometrische Diagnostizierbarkeit des Unrechtsbewußtseins bzw. des Vergeltungspflichtverständnisses mit Hilfe der quantitativen Urteile. Für den Verschulden-Effekt in der Instruktion (3. Gegenüberstellung) ergaben sich 1, 2, 2, 4 für die Vorschüler als kleinste Differenzen neben nur bei Vorschülern auftretenden 8 Null-Differenzen. Also waren noch 11 von 33 Vorschülern unterhalb des bei Erwachsenen erreichbaren Irrtumsniveaus. Die entsprechenden Werte waren 1, 1, 1, 2 für die Grundschüler (also 3 unterhalb der gewählten Erreichbarkeits-Grenze für die Diagnose des Unrechtsbewußtseins); 4, 5, 6, 8 für die Hauptschüler (also alle über der Grenze); 2, 4, 4, 4 für die Erwachsenen (also alle über der Grenze). Für den Entschuldigung-Effekt in der Instruktion (2. Gegenüberstellung) ergaben sich 1, 1, 3, 5 für die Vorschüler und 1 Null-Differenz. Also blieben hier nur 3 Vorschüler unterhalb der gewählten Grenze für das Vergeltungspflichtverständnis. Entsprechend 1, 6, 6, 7 für die Grundschüler (einer unter der Grenze); 2, 3, 5, 5 für die Hauptschüler (alle über der Grenze); 1, 2, 3, 4 für die Erwachsenen (alle über der Grenze). Wenn überhaupt ist die Diagnostizierbarkeit der beiden Einsichtskomponenten aufgrund der Instruktionsphase nur bei Grund- und Vorschülern ausgeblieben, bei letzteren häufiger und beim Unrechtsbewußtsein häufiger. Man wird aber berücksichtigen müssen, daß die Vorgabeform der Gegenüberstellung der Stimulusalternativen die 'Aufgabe erleichtert und nicht unbedingt ökologisch valide ist, da sie ja bei den Einsichtskomponenten nicht einbezogen sein muß. Das wäre aber auch durch den Zivilrichter zu entscheiden, so daß die Vornahme beider Meßmethoden in einer Begutachtung sinnvoll und berichtenswert wäre.

Die psychometrische Beurteilung der mittleren Effekte in der Instruktionsphase erscheint unter der Betrachtungsweise *de lege ferenda* in Hinsicht auf die expliziten und impliziten Annahmen des Zivilrechts zur Deliktsfähigkeit noch mitteilenswert, weil sich daraus mehr als die sonst in ihnen sichtbare Ausbildung der Alterstrends in den Effektstärken der drei Stimuluskomponenten wie in der Beurteilung dreifaktorieller Stimuli ergibt. Im Entschuldigung-Effekt wichen die mittleren Urteilsdifferenzen in der Instruktion psychometrisch beurteilt mit $t = 3.65, 9.19, 6.47$ und 6.61 (in der Altersreihenfolge) von der Null-Differenz ab. Dagegen ergab sich im Verschulden-Effekt, daß die Differenz der Vorschüler noch nicht einmal das 10%-Niveau überschritt. Denn die Statistiken betrugen $t = 1.40, 2.57, 4.83$ und 3.98 in der Altersreihenfolge. Demnach ist sowohl die Altersgrenze vom vollendeten siebenten Lebensjahr als auch die Indikationstheorie des Reichsgerichts durch Anlegen psychometrischer Maßstäbe besonders gestützt worden.

3.2.3 Fazit

Das Urteilsverhalten während der Instruktion und in der Hauptphase der Durchführung von Beurteilungen faktoriell zusammengesetzter Bild-Geschichten scheint demnach psychometrische Eigenschaften zu besitzen, die seine Verwendung in forensischen Untersuchungen zur Deliktsfähigkeit nach § 828 (2) BGB nahelegen würden. Darüber hinaus ließen sich auf der Grundlage der psychometrisch gesicherten Dia-

gnosen auch die impliziten zivilrechtlichen Annahmen zu § 828 BGB über den Schluß vom ausgebildeten Unrechtsbewußtsein auf das Vergeltungspflichtverständnis stützen.

3.3 Kovariationen mit Intelligenzmaßen

Von den 27 Interkorrelationen der drei Differenzmaße mit den drei wie zu erwarten untereinander korrelierenden Untertests des HAWIK-R für die drei Altersgruppen waren jedoch nur zwei auf dem 5%-Niveau zu sichern. Mit den Verschulden-Effekten korrelierten bei den Vorschülern der WT zu $r = .45$ und bei den Hauptschülern der MT zu $r = -.37$, wobei letzteres angesichts der mittleren Urteile und der Reliabilität der Effekte nur bedeutet, daß die im MT besseren Hauptschüler immer noch zuverlässig zwischen den Verschulden-Stufen unterschieden. Die Tabelle 5 gibt die Ergebnisse der multiplen Regression von Intelligenztestmaßen, Alter und Geschlecht auf die Urteile an. Die Zusammenhänge sind zwar insgesamt statistisch gesichert, jedoch diagnostisch wegen ihrer geringen systematischen Varianzanteile unbrauchbar für eine empirisch fundierte Ersetzbarkeit der drei Urteilsmaße durch die Prediktor-Variablen der IQ-Subtests, Alter und Geschlecht.

3.3.1 Fazit

Aus IQ-Maßen auf die Stärke der Berücksichtigung von Verschulden- oder Wiedergutmachung-Informationen in moralischen Urteilen zu schließen erscheint demnach kaum möglich.

Tabelle 4: a) Mittlere Urteile bzw. b) Urteilsdifferenzen (D) und deren Standardabweichungen (S) für vier Altersgruppen in der Instruktionsphase mit gegenüberstellenden Einzelbeurteilungen der Bedingungen der drei Stimulusfaktoren aus der Hauptphase (K. Ent.: Keine Entschuldigung; Ent.: Entschuldigung; Rep.: Reparaturhilfe).

a)	N	Hundehütte	Schuppen	K.Ent.	Ent.	Rep.	Versehen	Wut	
Vorschüler	33	-7.2	-9.5	-10.4	5.0	6.1	-3.5	-10.7	
Grundschüler	38	-8.0	-11.7	-10.3	5.0	8.8	-1.5	-11.5	
Hauptschüler	37	-5.8	-10.5	-9.5	6.2	9.5	1.2	-10.0	
Erwachsene	40	-4.7	-8.9	-7.5	4.4	9.2	-3	-9.5	
b)		D	S	D	S	D	S	D	S
Vorschüler	33	2.3	(8.7)	15.4	(9.0)	1.1	(7.0)	7.2	(7.8)
Grundschüler	38	3.7	(3.2)	15.3	(6.8)	3.8	(2.9)	10.0	(8.3)
Hauptschüler	37	4.7	(3.1)	15.7	(7.0)	3.3	(3.4)	11.2	(6.7)
Erwachsene	40	4.2	(2.7)	11.9	(6.0)	4.8	(3.6)	9.2	(5.3)

Tabelle 5: Ergebnis der Multiplen Regression als Overall-Test für die Einflüsse von MT-, WT- und AV-Punktwerten sowie von Alter und Geschlecht auf drei Effektmaße der Urteile von 108 Kindern.

	Individuelle Effekte von		
	Entschuldigung	Reparaturhilfe	Verschulden
Alter	-	.01	-
Geschlecht	.08	-	-
MT	.04	-	.03
WT	-	-	-
AV	-	-	-
R ²	.105	.198	.163
F	12.61	2.40	5.04
P	.000	.042	.000

Anmerkung:

Im oberen Bereich der Tabelle sind nur diejenigen p der Beta-Gewichte angegeben, die kleiner als .10 waren.

Vielmehr können sie nur die eigenständige Komponente des Wissen-Könnens erfassen. Die stellvertretende Verwendung von Intelligenzmaßen für die Kriterien der Deliktsfähigkeit im Sinne des Werten-Könnens kommt angesichts der fehlenden oder geringen Korrelationen, insbesondere mit AV, trotz hoher Reliabilität der Effektstärkenmaße entgegen der forensisch-psychiatrischen Auffassung Bressers (1972) oder der psychologischen Begutachtungspraxis (Ell, 1983, S. 65) nicht in Frage. Statt dessen wäre die eigenständige Erfassung sowohl der wertenden als auch der rein-intellektuellen (im Sinne des Wissen-Könnens) Einsichtsfähigkeiten forensisch erforderlich. Erstere zu erfassen erscheint mit den hier verwendeten Verfahren aus Gründen inhaltlicher Validität angesichts der Reliabilität sinnvoll, wobei allerdings dem Zivilrichter die letztgültige Entscheidung über das Vorliegen inhaltlicher Validität überlassen bleibt.

4. Diskussion

Die Untersuchung brachte Ergebnisse zu § 828 BGB hinsichtlich der beiden hilfswissenschaftlichen Aufträge der Psychologie gegenüber der Rechtspflege: *de lege ferenda* (Gesetzgebung) und *de lege lata* (Rechtsanwendung).

Zu den impliziten und expliziten Annahmen des Gesetzestexts (*de lege ferenda*) konnten mit dem spezifischen Delikt des Brandschadens kei-

nerlei Widersprüche gewonnen werden, wenn man in der moralischen Bewertungsunterscheidung der Entschuldigung die adäquate Operationalisierung der geringen Anforderung an das Vergeltungspflichtverständnis sieht, weiterhin den Anstieg in der individuellen Erfassbarkeit und die noch im Vorschulalter «disordinalen» Urteilsleistungen im Zusammenhang mit der Beurteilung des Reparierenhelfens und der Schadenshöhe als Stützungen der Ziehung der Altersgrenze auffaßt, deren geringe Anforderung sich in später auftretenden Differenzierungen bemerkbar macht. Insofern bestätigte sich das zusammenfassende Urteil von Hommers (1983) über die damals vorliegende nicht direkt auf die zivilrechtliche Altersgrenze abzielende empirische Ergebnislage der kognitiven Entwicklungspsychologie mit dem hier verwendeten spezifischen Delikt der Brandstiftung in einer direkt auf die erforderlichen kognitiven Inhalte bezogenen Untersuchung, die die relevanten Altersgruppen umfaßte.

De lege lata ist die auf die Begutachtung der zivilrechtlichen Verantwortlichkeit abzielende Feststellung «Zuverlässige, einhellige und begrifflich faßbare Kriterien gibt es nicht» (Eisen, 1977, S. 320) nicht mehr als zutreffend anzusehen, wenngleich damit nicht alle hier bestehenden Probleme der forensischen Begutachtung gelöst sind (z. B. das der auf den Tatzeitpunkt rückbezogenen Beurteilung, vgl. Wegener, 1981). Vielmehr ist gerade von hier aus interessant, daß die Reliabilitätsschätzungen der durch Wertungsdifferenzen erfaßten wertenden Komponenten des Vergeltungspflichtverständnisses und des Unrechtsbewußtseins gerade im Grundschulbereich hinreichend hoch entwickelt waren, wo die Begutachtungsaufträge im Falle von zündelnden Kindern anfallen. Die generelle, fast verständnisunwillig anmutende Kritik Bressers (1988) an einer statistische oder mathematisierend-experimentelle Methoden anwendenden Rechtspsychologie erscheint daher ungerechtfertigt (vgl. in diesem Sinne, aber zu einem anderen Punkt: Steller, 1989). Vielmehr kann und sollte gerade durch diese Methoden in der Tradition Marbes (1913; 1926, 6. Kap.) fortgeföhren werden, der sein Gutachten zum Müllheimer Eisenbahnunglück vom 17. Juli 1911 eben auf experimentelle und quantifizierende Methoden stützte.

Bei der Verwendung des vorgestellten Verfahrens in der forensisch-psychologischen Begutachtung (*de lege lata*) sollte man aber noch folgendes bedenken:

- Die Wahl des statistischen Niveaus muß hier, aber nicht überall in der Forensischen Psychologie, progressiv angesetzt werden, weil der Untersuchte einen Vorteil (nicht für den durch ihn verursachten Schaden haften zu müssen) zugesprochen bekommen soll, wenn er die Diagnose «Fehlende Einsicht» im Unrechtsbewußtsein oder im Vergeltungspflichtverständnis erfüllt.
- Natürlich gilt auch hier «Ein Test ist kein Test» als Leitlinie für die gutachterliche Beurteilung des Probanden. Man wird als Meta-Kriterium der Diagnose «Fehlende Einsicht» die Konsistenz der *multi-method*-Untersuchungsergebnisse ansehen müssen. Die vorgestellte Verfahrensweise hat sicherlich gegenüber explorativen Methoden Vorteile in der Angabe von Irrtumswahrscheinlichkeiten. Klare Äußerungen anderer Art eines Probanden sind deswegen aber noch nicht grundsätzlich zweitrangig. Forensische Begutachtung ist wie andere Begutachtung eine verschiedene Informationen integrierende psychodiagnostische Tätigkeit im Sinne von Wegener und Steller (1986) und Steller (1988).
- Schließlich wäre die individuelle Schätzung des Meßfehlers (direkter Ansatz nach Huber, 1973) ein sinnvolles Vorhaben, um die Unschärfen der gruppenstatistischen Schätzer aufzuheben. Das erfordert aber Meßwiederholung im Sinne eines zweiten Durchganges mit dem Probanden und muß einer weiteren Studie vorbehalten bleiben.

Zu wünschen wäre, den vorgestellten psychometrischen Ansatz auch auf andere kognitive rechtspsychologische Kriterien anzuwenden. Dazu wären entsprechende diagnostische Forschungsarbeiten, die die kognitive Entwicklungspsychologie mit der psychometrischen Einzelfalldiagnostik verbinden, in die Forensische Psychologie einzubringen. Hier ist unmittelbar an das Unrechtsbewußtsein für fahrlässige Brandschäden zu denken, wo sich der hier gewählte Urteilsansatz unmittelbar anwenden läßt. Aber auch für die Fragen der strafrechtlichen Verantwortlichkeit nach § 3 JGG wären

zumindest hinsichtlich des dort verlangten im Niveau höheren, in das System Kohlbergs (1963) einordbaren Unrechtsbewußtseins (vgl. Hommers, 1983, 1989) entsprechende diagnostische Forschungsarbeiten zu erbringen, die angesichts der von Colby, Kohlberg, Gibbs und Lieberman (1983) für die amerikanischen Untersuchungen berichteten hohen Reliabilitäten der Stufen des Kohlbergschen Systems vielversprechend erscheinen.

Literatur

- Achenbach, T. M. & Edelbrock, C. S. (1981). Behavioral problems and competencies reported by parents of normal and disturbed children aged four through sixteen. *Monographs of the Society for Research in Child Development* (Serial No. 188), Volume 46, No. 1.
- Bresser, P. H. (1972). Die Beurteilung der Jugendlichen und Heranwachsenden im Straf- und Zivilrecht. In H. Göppinger & H. Witter (Hrsg.), *Handbuch der Forensischen Psychiatrie. Band II*. Berlin: Springer.
- Bresser, P. H. (1988). Über die Grenzen psychiatrischer Dokumentation: Was wird nicht abgebildet? *Forensia*, 9, 163–173.
- Canter, D. (Ed.). (1980). *Fires and Human Behavior*. Chichester: Wiley.
- Colby, A., Kohlberg, L., Gibbs, J. & Lieberman, M. (1983). A longitudinal study of moral judgment. *Monographs of the Society for Research in Child Development* (Serial No. 200), Volume 48, No. 1–2.
- Dauner, I. (1980). *Brandstiftung durch Kinder. Kriminologische, kinderpsychiatrische und rechtliche Aspekte*. Bern: Huber.
- DeBoor, W. (1955). Zur Psychologie und Psychopathologie der Brandstiftung (1917–1955). *Fortschritte der Neurologie und Psychiatrie*, 23, 367–378.
- Eisen, G. (1977). *Handwörterbuch der Rechtsmedizin für Sachverständige und Juristen*. Stuttgart: Enke.
- Ell, E. (1983). *Wenn Kinder zünden, Vorschläge zur Feuererziehung*. Tübingen: Katzmann.
- Holland, C. J. (1969). Elimination of fire-setting in a seven-year-old boy. *Behaviour Research and Therapy*, 7, 135–137.
- Hommers, W. (1983). *Die Entwicklungspsychologie der Delikts- und Geschäftsfähigkeit*. Göttingen: Hogrefe.
- Hommers, W. (1986a). Non-Additivität als Beleg für die moralische Natur der Integration von Schaden und Ersatzleistungen. *Archiv für Psychologie*, 138, 71–90.
- Hommers, W. (1986b). Zusammenwirken von Schaden und Ersatzleistungen im moralischen Urteil. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 18, 12–21.
- Hommers, W. (1988a). Entschuldigung und Entschädigung für einen Diebstahl. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 20, 121–133.

- Hommers, W. (1988b). Die Wirkungen von Entschuldigung und Entschädigung auf Strafurteile über zwei Schadensarten. *Zeitschrift für Sozialpsychologie*, 19, 139–151.
- Hommers, W. (1989). Die Entwicklung der Einsicht in das Delikt. In S. Bäuerle (Hrsg.), *Kriminalität bei Schülern. Band 1. Ursachen und Umfeld von Schülerkriminalität*. Stuttgart: Angewandte Psychologie.
- Huber, H. P. (1973). *Psychometrische Einzelfalldiagnostik*. Weinheim: Beltz.
- Klosinski, G. (1985). Jugendliche Brandstifter und Sexualdelinquenten: Ein Vergleich der Psychopathologie, Familiensituation und Familiendynamik. *Forensia*, 5, 149–156.
- Kohlberg, L. (1963). The development of children's orientation towards a moral order: I. Sequence in the development of moral thought. *Vita Humana*, 6, 11–33.
- Marbe, K. (1913). Psychologische Gutachten zum Prozeß wegen des Mühlheimer Eisenbahnunglücks. *Fortschritte der Psychologie und ihrer Anwendungen*, 1, 339–374.
- Marbe, K. (1926). *Der Psycholog als Gerichtsgutachter im Straf- und Zivilprozess*. Stuttgart: Enke.
- Prins, H. (1986). *Dangerous Behavior, the Law, and Mental Disorder*. London: Tavistock.
- Steller, M. (1988). Standards der forensisch-psychologischen Begutachtung. *Monatsschrift für Kriminologie und Strafrechtsreform*, 71, 16–27.
- Steller, M. (1989). Gemälde oder Bauwerke? Anmerkungen zur Objektivierung der forensischen Begutachtung. *Monatsschrift für Kriminologie und Strafrechtsreform*, 72, 155–159.
- Tewes, U. (1983). *Hamburg-Wechsler-Intelligenztest für Kinder, Revision 1983*. Bern: Huber.
- Undeutsch, U. (1967). Delikthaftung junger Menschen. In U. Undeutsch (Hrsg.), *Forensische Psychologie* (S. 567–597). Göttingen: Hogrefe.
- Walster, E., Walster, G. W. & Berscheid, E. (1978). *Equity: Theory and Research*. Boston: Allyn and Bacon.
- Wegener, H. (1981). *Einführung in die Forensische Psychologie*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Wegener, H. & Steller, M. (1986). Psychologische Diagnostik vor Gericht. Methodische und ethische Probleme forensisch-psychologischer Diagnostik. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 7, 103–126.

Anschrift: Prof. Dr. Wilfried Hommers, Institut für Psychologie (I) der Universität Würzburg, Domerschulstraße 13, 8700 Würzburg

1. Einleitung

Eines der am häufigsten gefundenen Ergebnisse der schulischen Leistungsangst-Forschung ist die Tatsache, daß Mädchen generell höhere Angstwerte erzielen als Jungen (z. B. Nickel & Schlier, 1970; Garnet-Harnack, 1972; Koppes, 1975; Jacobs & Steinmutter, 1979). Fast alle Autoren haben jedoch Bedenken gegen eine kritische Interpretation dieses Resultats. So äußerten bereits Sarason, Davidson, Lightshill, Wade und Radash (1971), daß es sich hierbei nicht um reale Angstunterschiede, als vielmehr um die Folge von geschlechtsspezifischer Er-

ziehung handle, welche es Mädchen eher erlaube, ihre Angst zu zeigen. Das gefundene Geschlechterunterschiede wurden auch darauf zurückgeführt, daß die herkömmlichen Angstfragebögen zunehmen teilweise geschlechtsspezifische Items enthalten (Schwarzer, 1975). Darüber hinaus sind bei fast allen bekannten Untersuchungen zur Leistungsangst sehr große Stichproben verwendet worden, wodurch bereits geringfügige und praktisch nachweisbar wenig relevante Abweichungen in den Angstwerten statistisch signifikant werden konnten. Es stellt sich somit die Frage, inwieweit die gefundenen Geschlechterdifferenzen evtl. eine Angstunter-